Accession No. 24049

Copy No.

SID 63-741

AN ANALYSIS OF TWO-IMPULSE

ORBITAL TRANSFER

July 1, 1963

Prepared by

Gentry Lee

SPACE SCIENCES LABORATORY

Approved by

E. R. van Driest
Director

NORTH AMERICAN AVIATION, INC.

SPACE and INFORMATION SYSTEMS DIVISION

FOREWORD

The work presented in this document was performed for the Marshall Space Flight Center under contract NAS 8-5211 (Satellite Rendezvous).

ACKNOWLEDGMENTS

To Mr. H. W. Bell, the assistant director of the Space Sciences Laboratory, the author extends his gratitude--both for his patient advice during the research and the vital technical contributions that provided the stimuli for much of this work. I also am indebted to Dr. D. F. Bender, Mr. G. A. McCue, and the entire Space Mechanics section of the laboratory; their prior research in this area laid the foundation for this work.

ABSTRACT

Analytical investigations of two-impulse transfers between elliptical orbits, using vector analysis and other mathematical techniques, have yielded pertinent, heretofore unknown facts about an orbital transfer function. One particular mode of analysis, the Bell-Arenstorf technique, helped show not only that the minimum velocity increment solution between two points on elliptical orbits could be along a hyperbola, but also that there could be two relative minima in this impulse function. Particular examples of both these phenomena are given.

Although prior published analysis has been restricted mostly to coplanar elliptical orbits, this analysis includes inclined elliptical orbits. An eighth-order polynomial expression, the real roots of which may refer to extrema in the impulse function, is determined. Since it can be shown that some of these roots are extraneous--not corresponding to impulse minima--two test functions are next determined that define regions in which all extrema must lie. These regions identify those roots that do correspond to extrema in the impulse function and those that are extraneous. These new analytical findings have been incorporated into an earlier computer contour mapping program that locates the optimum transfer between elliptical orbits.

TABLE OF CONTENTS

LIST OF FIGURES

NOMENCLATURE

Scalars

a	Semi-major axis
e	Eccentricity
i	Inclination
Ω	Right ascension of ascending node
ω	Argument of perigee, angle from reference axis to perigee point
p	Semi-latus rectum
$\Delta\theta$	True anomaly angle traversed in transfer orbit plane
μ	Gravitation constant
ϕ_1	Angle from reference axis to departure position in initial orbit.
ϕ_2	Angle from reference axis to arrival position in terminal orbit
α	Angle between $\underline{r}_2$ and $\underline{r}_2 - \underline{r}_1$
β	Angle between $\underline{r}_1$ and $\underline{r}_1 - \underline{r}_2$
z	Term defined in equation (14)
Ψ_1	Functional form of 1st velocity increment
Ψ_2	Functional form of 2nd velocity increment
I	Functional representation of impulse
I^*	Function defined by equation (41), whose extrema are also located in equation (37)
f	One of test functions used in analyzing short transfer

<u>Scalars</u>

g	The other test function used in analyzing short transfer
h	One of test functions used in analyzing long transfer
k	Other test function used in analyzing long transfer
A-H	Coefficients that determine interval-finding polynomials
ϕ_1-ϕ_9	Coefficients that determine minimizing polynomial
σ	Function defined by equation (51)
τ	Function defined by equation (52)

<u>Vectors</u>

$\underline{e}$	Orbit shape and orientation vector
$\underline{r}_1$	Vector from reference position to point of departure on initial orbit
$\underline{r}_2$	Vector from reference position to point of arrival on final orbit
$\underline{W}$	Unit vector directed along orbit's angular momentum vector
$\underline{v}$	Vector defined by equation (13)
$\underline{V}_{tj}$	Velocity vectors in transfer orbit
$\underline{V}_j$	Velocity vectors in initial and final orbit
$\underline{V}_{par}$	Velocity vector in parabolic orbit
$\underline{V}_r$	Velocity vector in circular orbit
$\underline{U}_j$	Unit vectors in direction of radius vectors
$\underline{W},\underline{I},\underline{W}_2$	Unit vectors in Cartesian coordinates defining the reference plane

I. INTRODUCTION

One of the major problems of the nascent space age is concerned with changing orbits in space. To transfer from orbit to orbit can require immense quantities of fuel, far beyond the limitations of today's engineering. It is, therefore, of extreme practical interest to be able to locate particular modes of transfer between these orbits that use the least possible fuel.

The most general problem of optimum two-impulse orbital transfer, in which the chief assumption is that the elliptical orbits are unperturbed, permits both the departure point and the arrival point to be arbitrary and finds the single best mode of transfer between the two given orbits. The most general constraint is to fix the end points; then the optimization procedure is carried out solely along a parameter defining all the transfer orbits that go through these given terminals.

The impulsive case of orbital transfer is, of course, an ideal situation. There is one instantaneous thrust from the initial orbit into the transfer orbit; there is a second instantaneous thrust to get into the final orbit. The information gained from the solution of this problem should provide a basis for the study of orbital transfer with finite thrust.

The majority of the published two-impulse orbital transfer work deals with fixed terminals and co-planar orbits(2-4). The recent work by G. A. McCue(5) represents an extension, by means of a numerical contour mapping approach, to include both free terminals and inclined orbits. The analysis in the paper presented here is valid for inclined orbits and is directed towards solving the fixed terminal problem. However, since it is necessary to solve this problem many times before the free terminal problem can be investigated by means of contour maps, the findings of this analysis have been incorporated into Mr. McCue's numerical program.

II. EXPLANATION OF PROBLEM FORMULATION

Two Keplerian elliptical orbits in space can be defined by their orbital elements, a, e, i, Ω, and ω. In the general two-impulse orbital transfer problem, it is desirable to locate the minimum velocity increment solution between any two such Keplerian orbits. If the plane of the second orbit of the transfer is the reference plane, then i_2, the inclination of the second orbit, is zero. The terms ordinarily referred to as the "nodal" parameters, Ω_1, and Ω_2, are made zero by selecting the line of intersection of the two orbit planes as the reference direction. (See Figure 1)

$$\left(\underline{N} = \frac{\underline{W}_2 \times \underline{W}_1}{|\underline{W}_2 \times \underline{W}_1|} \right) \tag{1}$$

This leaves seven orbital elements (a_1, e_1, i_1, ω_1 and a_2, e_2, ω_2--subscripts one and two refer to elements in the first and second orbits respectively) that define the two orbits between which the transfer is to be accomplished.

Three variables which define all possible means of transferring from the first orbit to the second orbit are ϕ_1, the angle from reference line ($\underline{N}$) to a departure point on the first orbit; ϕ_2, the angle from reference line to arrival point on second orbit; and p, the semi-latus rectum of the transfer orbit between the two points. The parameter p is chosen as the third variable because it simplifies the nature of the impulse function. Other formulations for the third variable can produce serious discontinuities [5].

The "total impulse" used in transferring between the orbits is defined as the sum of the magnitudes of the velocity changes necessary to get from the first orbit into the transfer orbit and then from the transfer orbit into the second orbit. In this paper, an <u>optimum</u> impulse solution refers to a particular configuration of the <u>three</u> variables that leads to the least possible impulse between <u>two orbits</u>. A <u>minimum</u> impulse solution refers to the transfer orbit which gives the least total impulse for a <u>given</u> arrival-point, departure-point configuration.

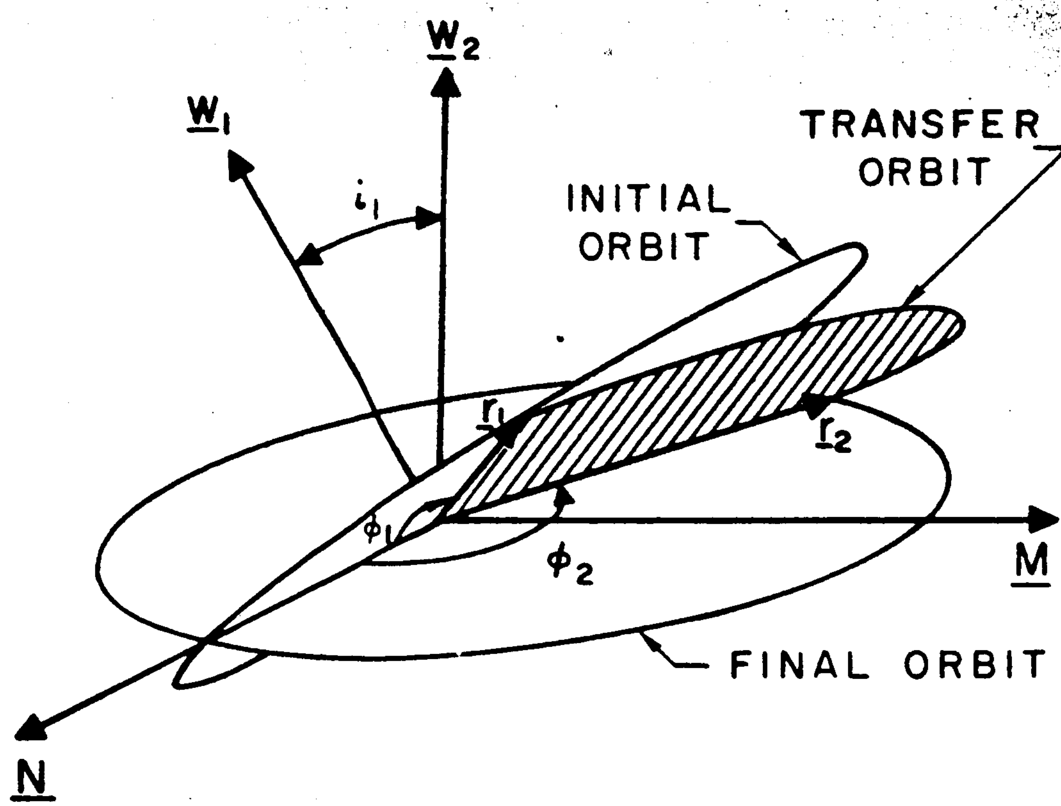


FIGURE 1. TRANSFER GEOMETRY

III. TRANSFER GEOMETRY

In rendering the orbital transfer problem subject to analysis, it is most convenient to express the important quantities in their vector representation. The vectors $\underline{r}_1$ and $\underline{r}_2$ represent the vectors from the attracting body to the departure and arrival points. Define unit vectors $\underline{U}_1$ and $\underline{U}_2$ in the direction of $\underline{r}_1$ and $\underline{r}_2$. The components of these vectors are then given by

$$\underline{U}_1 = \cos \phi_1 \, \underline{i} + \sin \phi_1 \cos i_1 \, \underline{j} + \sin \phi_1 \sin i_1 \, \underline{k} \tag{2}$$

$$\underline{U}_2 = \cos \phi_2 \, \underline{i} + \sin \phi_2 \, \underline{j} \tag{3}$$

$$\underline{r}_m = \left[\frac{p_m}{1 + e_m \cos(\phi_m - \omega_m)} \right] \underline{U}_m, \quad m = 1,2 \tag{4}$$

where $\underline{i}$, $\underline{j}$, and $\underline{k}$ are unit vectors in a right-handed Cartesian system with $\underline{i}$ in the direction of $\underline{N}$.

Three more useful vectors in the analysis are $\underline{W}_1$, $\underline{W}_2$, and $\underline{W}_t$— these are normal to the initial, final, and transfer orbit planes and are defined as follows:

$$\underline{W}_1 = -\sin i_1 \; \underline{j} + \cos i_1 \; \underline{k} \tag{5}$$

$$\underline{W}_2 = \underline{k} \tag{6}$$

$$\underline{W}_t = \frac{\underline{U}_1 \times \underline{U}_2}{|\underline{U}_1 \times \underline{U}_2|} \quad \text{where } |\underline{U}_1 \times \underline{U}_2| \neq 0 \tag{7}$$

To complete the vector description, define two vectors $\underline{e}_1$ and e_2— these define the shape and orientation of the two orbits[6].

$$\underline{e}_m = e_m\ (\cos\omega_m\ \underline{i} + \sin\omega_m \cos i_m\ \underline{j} + \sin\omega_m \sin i_m\ \underline{k}) \tag{8}$$

where $m = 1,2$

The part of the transfer orbit traversed in the transfer is a certain true anomaly interval $\Delta\theta$. This interval may be quickly determined from

$$\cos\Delta\theta = (\underline{U}_1\ .\ \underline{U}_2) \quad 0° < \Delta\theta < 180° \tag{9}$$

No generality is lost if the true anomaly interval is limited to the first two quadrants. Although this does restrict the problem to "short transfers," if the signs of the velocity vectors in the transfer orbit are changed, the long transfers may be considered. The singularities in the impulse function at $\Delta\theta = 180°$ and $\Delta\theta = 0°$ indicate that the problem is simplified by considering the long and short transfers separately. Thus, in order to determine the absolute optimum transfer between two elliptical orbits, it is necessary to compare the optima found from all the short transfers and all the long transfers.

For every elliptical transfer orbit between a given departure point and arrival point, there exists both a short transfer and a long transfer. However, when considering particular hyperbolic transfer orbits, it is important to realize that either the short transfer or the long transfer is meaningless—it would require going out to infinity and back.

IV. THE BELL-ARENSTORF TECHNIQUE

The "Bell-Arenstorf technique" refers to a geometrical method of analyzing the two impulse orbital transfer problem. This methcd is based upon some cogent variable relationships recognized separately by Mr. H. W. Bell of North American Aviation, Inc.,[7] and Dr. Richard Arenstorf of Marshall Space Flight Center[8]. The fundamental idea of the method - that all possible transfers between fixed terminals on any elliptical orbits can be represented by two hyperbolae - provided the stimulus for much of this analysis.

For any two elliptical orbits, let $\underline{r}_1$ and $\underline{r}_2$ be the vectors from the reference position on the line of intersection to the departure and arrival points, respectively. The angle between them is $\Delta\theta$ and the size of this angle can be selected to be always in the first two quadrants without any loss of generality. By forming the vector $\underline{r}_2 - \underline{r}_1$, a triangle is made of the three vectors in the transfer orbit plane.

Define the two angles α and β (Figure 2) as follows:

$$\beta = \arcsin \frac{|\underline{r}_2| \sin \Delta\theta}{|\underline{r}_2 - \underline{r}_1|} \tag{10}$$

$$\alpha = \pi - (\beta + \Delta\theta) \tag{11}$$

Consider the locus of all possible velocity vectors that can act upon the point defined by $\underline{r}_1$ and trace a conical orbit path that goes through the point defined by $\underline{r}_2$. This locus defines all possible conic transfer orbits between the two points, since a particular orbit is uniquely defined by its velocity vector at a given position.

The velocity vector of any transfer orbit at the particular point $\underline{r}_1$ is given by (See Appendix 1)

$$\underline{V}_{t1} = \underline{v} + z\, \underline{U}_1 \tag{12}$$

where

$$\underline{v} = \frac{(\mu p)^{\frac{1}{2}} \ (\underline{r}_2 - \underline{r}_1)}{|\underline{r}_1 \times \underline{r}_2|} \tag{13}$$

$$z = \left(\frac{\mu}{p}\right)^{\frac{1}{2}} \tan \frac{\Delta\theta}{2} \tag{14}$$

where p is the semi-latus rectum of the transfer orbit. Then V_{t1} may be written as a function of this variable p.

$$V_{t1}(p) = \left(\frac{\mu}{p}\right)^{\frac{1}{2}} \left[\frac{p\ |\underline{r}_2 - \underline{r}_1|}{|\underline{r}_1 \times \underline{r}_2|} \underline{m} + \tan \frac{\Delta\theta}{2}\ \underline{U}_1 \right] \tag{15}$$

where $\underline{m}$ is unit vector in the direction of $\underline{r}_2 - \underline{r}_1$. Every positive value of p greater than zero defines a certain transfer orbit whose velocity vector at $\underline{r}_1$ has components in the direction of $\underline{m}$ and $\underline{U}_1$.

For any coordinate axes, the locus of all points such that the product of the coordinates is a constant forms a hyperbola with the axes as asymptotes. Since the product of the magnitudes of the components in the $\underline{m}$-direction and the $\underline{U}_1$ direction is independent of p,

$$\left[\frac{(\mu p)^{\frac{1}{2}}\ |\underline{r}_2 - \underline{r}_1|}{|\underline{r}_1 \times \underline{r}_2|} \right] \left[\left(\frac{\mu}{p}\right)^{\frac{1}{2}} \tan \frac{\Delta\theta}{2} \right] = \frac{\mu \tan \frac{\Delta\theta}{2}\ |\underline{r}_2 - \underline{r}_1|}{|r_1 \times r_2|} \tag{16}$$

the formulation of V_{t1} defines a hyperbola with the oblique coordinates established by $\underline{m}$ and $\underline{U}_1$ as asymptotes. Thus the locus of all possible velocity vectors leaving $\underline{r}_1$ and arriving at $\underline{r}_2$ on a conic path forms a hyperbola.

Similarly, at r_2 the velocity vector for any transfer orbit (dependent upon its semi-latus rectum) is given by

$$\underline{V}_{t2} = \underline{v} - z\,\underline{U}_2 \tag{17}$$

This defines another hyperbola that represents the locus of all possible transfer orbits leaving from $\underline{r}_1$ and arriving at $\underline{r}_2$. These are shown in Figure 2. It is important to note that for <u>every</u> p, there is one point on <u>each</u> of these hyperbolae that represents the transfer orbit.

These two hyperbolae refer to the so-called short transfer, in which the true anomaly interval traversed in the transfer orbit is less than 180°. If the true anomaly interval is greater than 180° ("long transfer"), the other branches of these same two hyperbolae represent the locus of all transfer orbits. These are obtained by simply changing the sign of $\underline{V}_{t1}$ and $\underline{V}_{t2}$.

In Figure 2, the vectors $\underline{V}_1$ and $\underline{V}_2$, defining the initial and final orbits, are in the transfer orbit plane to simplify the analysis. Then this <u>particular</u> Bell-Arenstorf diagram represents a coplanar transfer and $\underline{V}_1$ and $\underline{V}_2$, defined by

$$\underline{V}_1 = \left(\frac{\mu}{p_1}\right)^{\frac{1}{2}} \underline{W}_1 \times (\underline{e}_1 + \underline{U}_1) \tag{18}$$

$$\underline{V}_2 = \left(\frac{\mu}{p_2}\right)^{\frac{1}{2}} \underline{W}_2 \times (\underline{e}_2 + \underline{U}_2) \tag{19}$$

must have magnitudes less than parabolic speed ($\underline{V}_{par}$).

$$\underline{V}^2_{par} = \frac{2\mu}{r} \tag{20}$$

In the Bell-Arenstorf diagram, the vectors $\underline{V}_1$ and $\underline{V}_2$ (which uniquely define the initial and final orbits) emanate from $\underline{r}_1$ and $\underline{r}_2$ and must lie within a certain radius containing all elliptical orbits.

In finding the minimum velocity change solution for this two-impulse case, the function to be minimized is

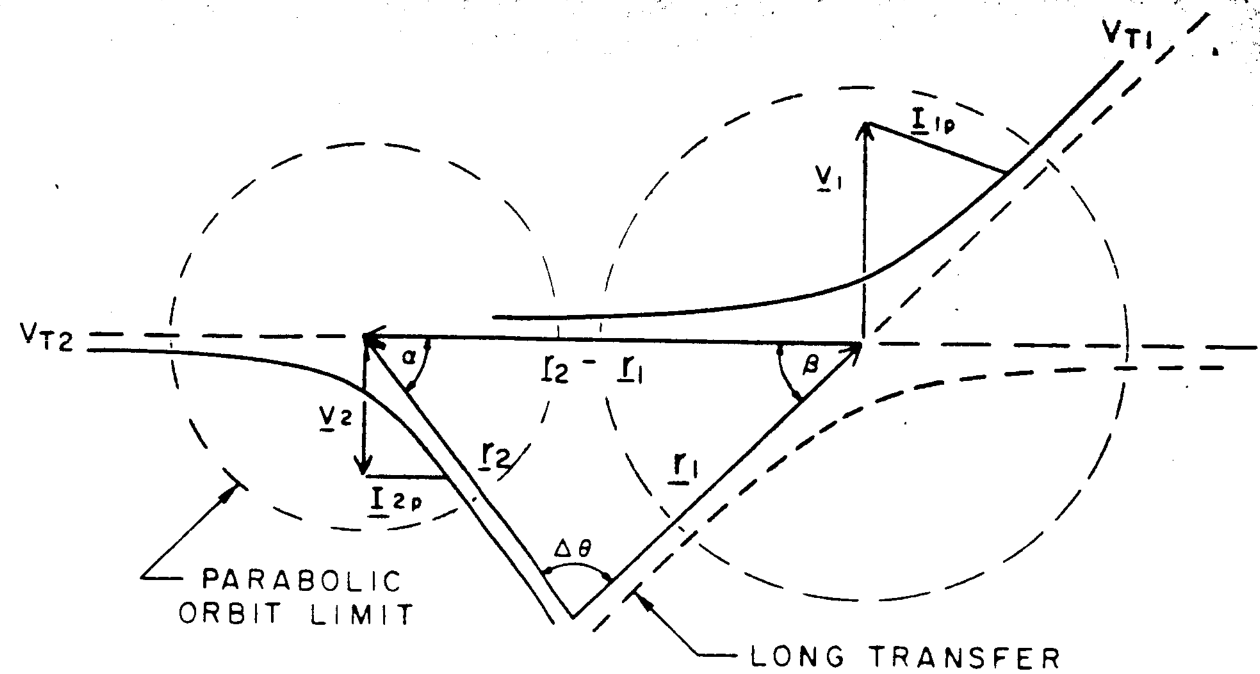
$$I(p) = \Psi_1(p) + \Psi_2(p) \tag{21}$$

where

$$\Psi_1(p) = | \pm \underline{V}_{t1}\ (p) - \underline{V}_1 | \tag{22}$$

$$\Psi_2(p) = | \underline{V}_{2-} \mp V_{t2}\ (p) | \tag{23}$$

The double sign on the transfer velocity vector refers to short and long transfers (upper sign is short). In the diagram, this optimization procedure requires that the sum of the distances from $\underline{V}_1$ and $\underline{V}_2$ to their respective transfer loci be minimized. For every p, there is one and only one point on each hyperbola corresponding to that transfer orbit. The distances marked I_{1p} and I_{2p} (In Figure 2) represent simply a particular transfer orbit chosen for illustrative purposes. The sums of their magnitudes would represent the impulse necessary to transfer between these two points along that particular conic.



$$V_{T1} = \left(\frac{\mu}{p}\right)^{1/2} \left[\frac{p(\underline{r}_2 - \underline{r}_1)}{|\underline{r}_1 \times \underline{r}_2|} + \mathrm{TAN}\ \frac{\Delta\theta}{2}\ \underline{U}_1\right]$$

$$V_{T2} = \left(\frac{\mu}{p}\right)^{1/2} \left[\frac{p(\underline{r}_2 - \underline{r}_1)}{|\underline{r}_1 \times \underline{r}_2|} - \mathrm{TAN}\ \frac{\Delta\theta}{2}\ \underline{U}_2\right]$$

FIGURE 2. BELL-ARENSTORF TECHNIQUE

V. APPLICATION OF BELL-ARENSTORF TECHNIQUE

The Bell-Arenstorf technique provides an excellent geometrical image of what is occurring in the two-impulse orbital transfer. By comparing the magnitudes of the impulse vectors for different transfer orbits, one can gain an intuitive feeling for the size of the impulse for a particular transfer orbit. More important, though, was the fact that the Bell-Arenstorf technique offered clues to two of the more important questions in the field.

In Mr. McCue's paper [5] he conducts a numerical search for the minimum impulse for each arrival-point, departure-point configuration and then, by a method of contour mapping, locates the optimum transfer between any two elliptical orbits. One of his early assumptions was that there could only be one minimum in the impulse function (variable p, semi-latus rectum of transfer orbit) for a fixed pair of terminals. The Bell-Arenstorf technique clearly showed the existence of a double minimum for a certain case and thus implied the existence of certain configurations under which a double minimum may be present.

It has been implied in nearly all of the definitive analytical works in this area, such as that by Altman[3], that the minimum velocity increment solution between points on elliptical orbits was always an ellipse. The Bell-Arenstorf technique suggested the existence of hyperbolic minima for certain configurations—this fact was subsequently proved.

The use of the Bell-Arenstorf technique stimulated further analytic investigations whose findings have been incorporated into Mr. McCue's optimization program.

VI. LOCATION OF DOUBLE MINIMUM

In order to assert that there can be a double minimum in the impulse function for fixed terminals, it is necessary only to find an example. By considering a particular case with unique symmetry properties, this example can be readily illustrated.

Consider the case where $|r_1| = |r_2|$ (Figure 3). This makes the angle α (Figure 2) equal to the angle β. Then the hyperbolae formed between the oblique axes at both the departure point and the arrival point are equivalent. If the entire coordinate system at $\underline{r}_2$ were flipped over and translated to $\underline{r}_1$, then these two hyperbolae would become coincident— they would match up point for point, transfer orbit for transfer orbit. Then the impulse function for particular elliptical orbits (defined by $\underline{V}_1$ and $\underline{V}_2$, both of which now act at the same point) is only the sum of the distances from $\underline{V}_1$ and $\underline{V}_2$ to all points on the hyperbola. Then, for this case, the minimum impulse solution corresponds to the point on the hyperbola from which the sum of the distances to $\underline{V}_1$ and $\underline{V}_2$ is a minimum.

For points with equal radii, one possible transfer orbit corresponds to a circular transfer. This transfer has a velocity vector ($\underline{V}_r$) perpendicular to the radius vector and its magnitude is given by

$$\underline{V}_r^2 = \frac{\mu}{r} \tag{22}$$

All velocity vectors emanating from $\underline{r}_1$ that have magnitudes less than $\left(\frac{2\mu}{r}\right)^{\frac{1}{2}}$ define elliptical initial and final orbits. In the diagram this range for $\underline{V}_1$ and $\underline{V}_2$ is described by a circle marked parabolic orbit limit.

Suppose $\underline{V}_1$ and $\underline{V}_2$ are located in such positions (See Figure 3), relative to each other, that the line connecting them intersects the hyperbola (either short transfer branch or long transfer branch) twice. As p varies from zero to its unbounded upper value, all possible transfer orbits have a corresponding point on the hyperbola. As p increases along the hyperbola, the value of the impulse is obviously decreasing until p reaches the value corresponding to a, where the line between $\underline{V}_1$ and $\underline{V}_2$ intersects the hyperbola. For values of p slightly larger than a (such as the p corresponding to point b), according to the triangle inequality the impulse must be higher. Thus the value of p at a must constitute a relative minimum in the impulse function.

As p nears the value corresponding to point c on the hyperbola, the triangle inequality states that the necessary transfer impulse is going down again. For points past c, the impulse is rising again, and thus c must also be a relative minimum. The fact that there can be two minima is thus demonstrated.

Numbers were placed into the diagram and indeed a double minimum (See Figure 4) occurred. The orbital elements for that particular fixed terminal case are given on the graph. For this case, the long transfer provides a greater impulse requirement for all transfers--thus only the short transfer is plotted.

In Appendix 2 a short mathematical investigation of the criteria for the existence of the double minimum in the case of equal radii is carried out. This investigation, which did lend some intuitive understanding to the problem, was not easily extendable to the case of non-equal radii.

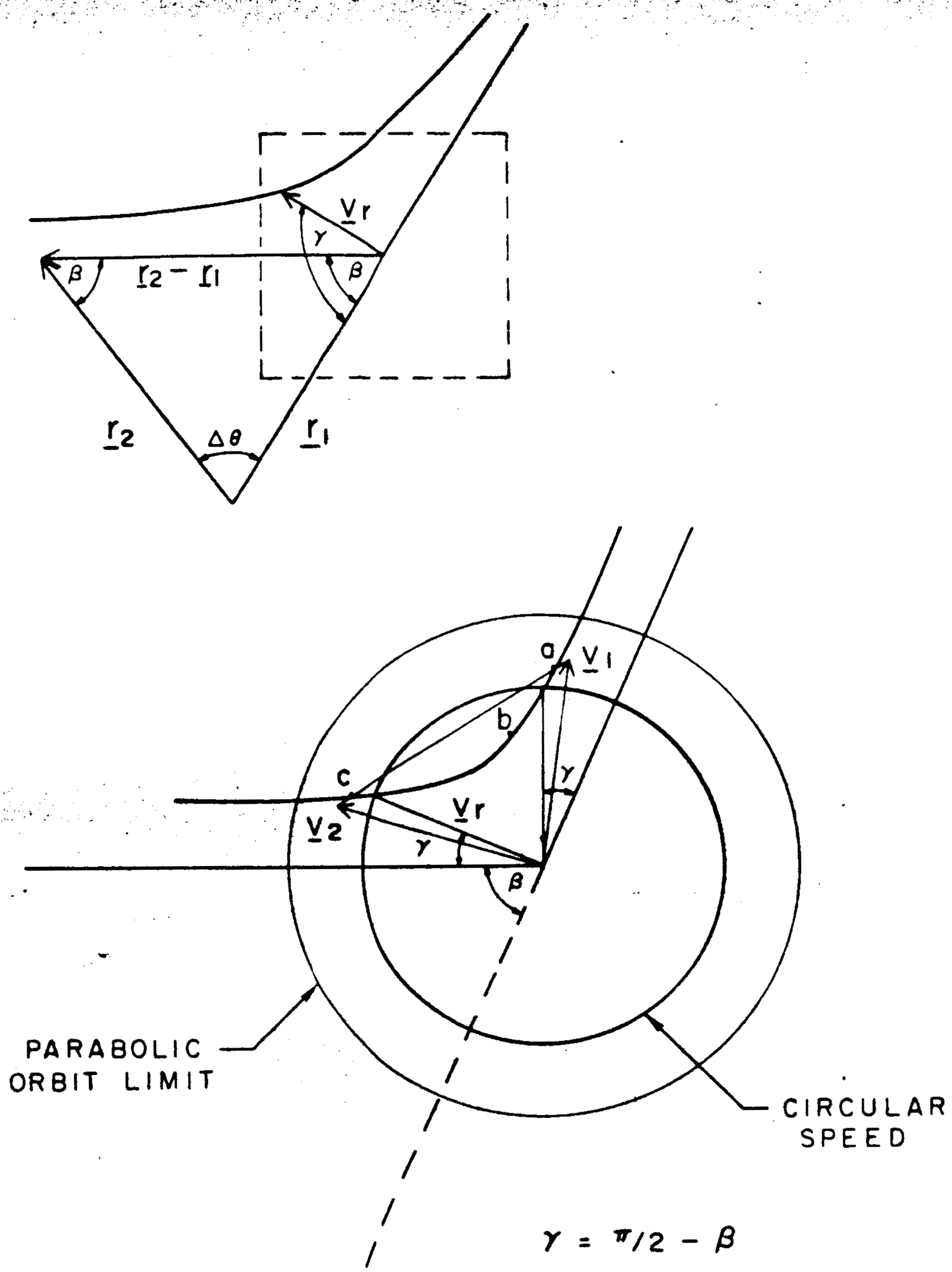


FIGURE 3 DOUBLE MINIMUM CASE OF EQUAL RADII

INITIAL ORBIT		FINAL ORBIT
6582	a (miles)	6582
0.93725	e	0.28355
0.0	i (degrees)	0.0
0.0	ω (degrees)	171.47

$\phi_1 = 153.7°$ $\phi_2 = 213.7°$

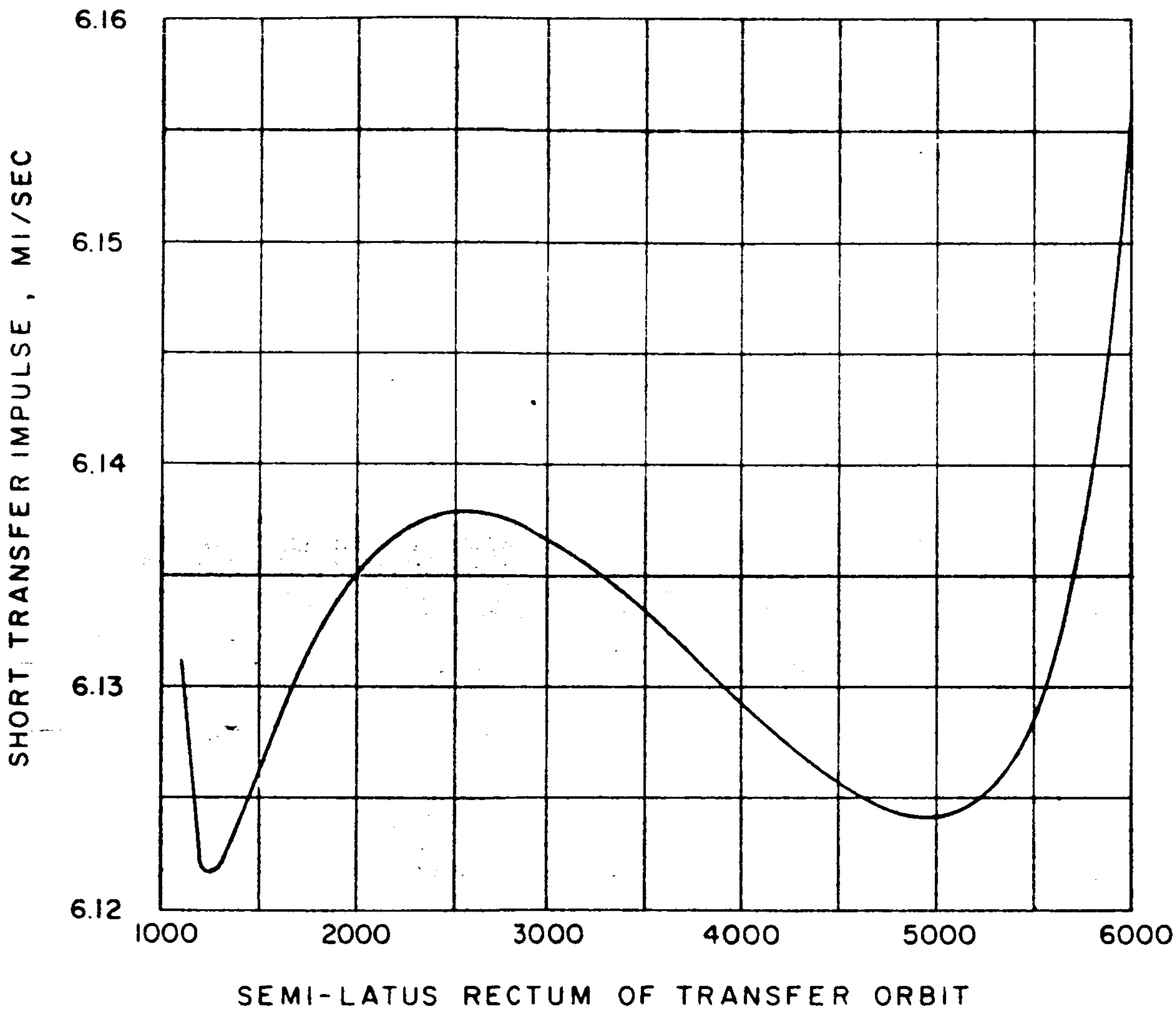


FIGURE 4. EXAMPLE OF DOUBLE MINIMUM

VII. LOCATION OF HYPERBOLIC MINIMUM

The assumption has been made, in prior two-impulse orbital transfer studies, that the minimum transfer between two points on elliptical orbits always lies along an ellipse. Although this has never been proved, it has been generally accepted. Use of the Bell-Arenstorf technique showed this assumption to be false.

For the case of $|\underline{r}_1| = |\underline{r}_2|$, it is clear from Figure 5 that a hyperbolic minimum may exist. Once again, the coordinate system at $\underline{r}_2$ is rotated and flipped such that all possible transfer orbits are given by one hyperbola. If the vectors $\underline{V}_1$ and $\underline{V}_2$ lie in the shaded region (see insert), the shortest distance from each to the transfer orbit hyperbola arrives at a point on that hyperbola outside the parabolic orbit limit. Since the least velocity increment - both to arrive in the transfer orbit and depart from it - lies along hyperbolic transfers, the sum of the two, the impulse, must have its minimum along a hyperbolic orbit between these two.

In investigating the more general case of non-equal radii, the geometry yielded not only configurations for which the minimum velocity increment solution could lie along a hyperbola, but also some other interesting properties about this orbital transfer function.

The general Bell-Arenstorf technique diagram can be modified (See Figure 6) in such a way as to orient both hyperbolae about the same coordinate axis with a common asymptote. Since for every p, the semi-latus rectum of the transfer orbit, there is one and only one point on each hyperbola, some manner of relating corresponding transfer orbits must be found. In Appendix 3 it is shown that there exist two families of circles, with centers on the $\underline{r}_2 - \underline{r}_1$ axis and radii dependent on the parameter p, that intersect the hyperbolae in such a way as to identify the points referring to the same transfer orbit. One family refers to the short transfers; the other, to the long. The short transfer family begins at the origin, with a member of infinite radius, and moves left through all possible values of p; the long transfer family goes in the opposite direction, also with increasing radius magnitude.

For every fixed arrival-point, departure-point configuration, there are two bounds on the values of the semi-latus rectum of the transfer orbit that define all elliptical transfer orbits. These "parabolic orbit limits" are defined by [5]

$$p_{max} = \frac{r_1 r_2 - \underline{r}_1 \cdot \underline{r}_2}{r_1 + r_2 - (2r_1 r_2 + 2\underline{r}_1 \cdot \underline{r}_2)^{\frac{1}{2}}} \tag{24}$$

$$p_{min} = \frac{r_1 r_2 - \underline{r}_1 \cdot \underline{r}_2}{r_1 + r_2 + (2r_1 r_2 + 2\underline{r}_1 \cdot \underline{r}_2)^{\frac{1}{2}}} \tag{25}$$

In the Bell-Arenstorf diagram, as p increases from zero to infinity, the radii of the family of circles diminish for both the long and short transfers. It is important to note that for $p > p_{max}$, the long transfer's being along a hyperbola is meaningless; similarly, for $p < p_{min}$, the short transfer implies going out to infinity to complete the orbit.

Regardless of what $\underline{r}_1$ and $\underline{r}_2$ are, there exists some value of p that defines the lower limit of elliptical transfer orbits. The circle marked "parabolic orbit limit" has its center at a point that is a value of p at which the long transfers change from hyperbolic into elliptical. Even though it is true that for every $\underline{r}_1$ and $\underline{r}_2$ this circle is located at a different place, it is important that it does exist somewhere and thus can be located arbitrarily. Then all initial orbits whose velocity vector at $\underline{r}_1$ lies inside the circle of radius A are elliptical; similarly for all final orbits whose velocity vector at $\underline{r}_2$ lies inside the circle of radius B. Suppose the initial and final orbits define velocity vectors $\underline{V}_1$ and $\underline{V}_2$ such that they are located as in Figure 5. It is clear that the least first increment change (to get into the transfer orbit) and the least final increment change are to transfer orbits that are hyperbolic. It is an easy extension to see that the sum of these two is a minimum along a hyperbola somewhere between these.

In Figure 7, impulse is plotted against the semi-latus rectum of the transfer orbit for a particular configuration. The parabolic orbit limits are marked and the orbital parameters are given--clearly the minimum transfer is along a hyperbola.

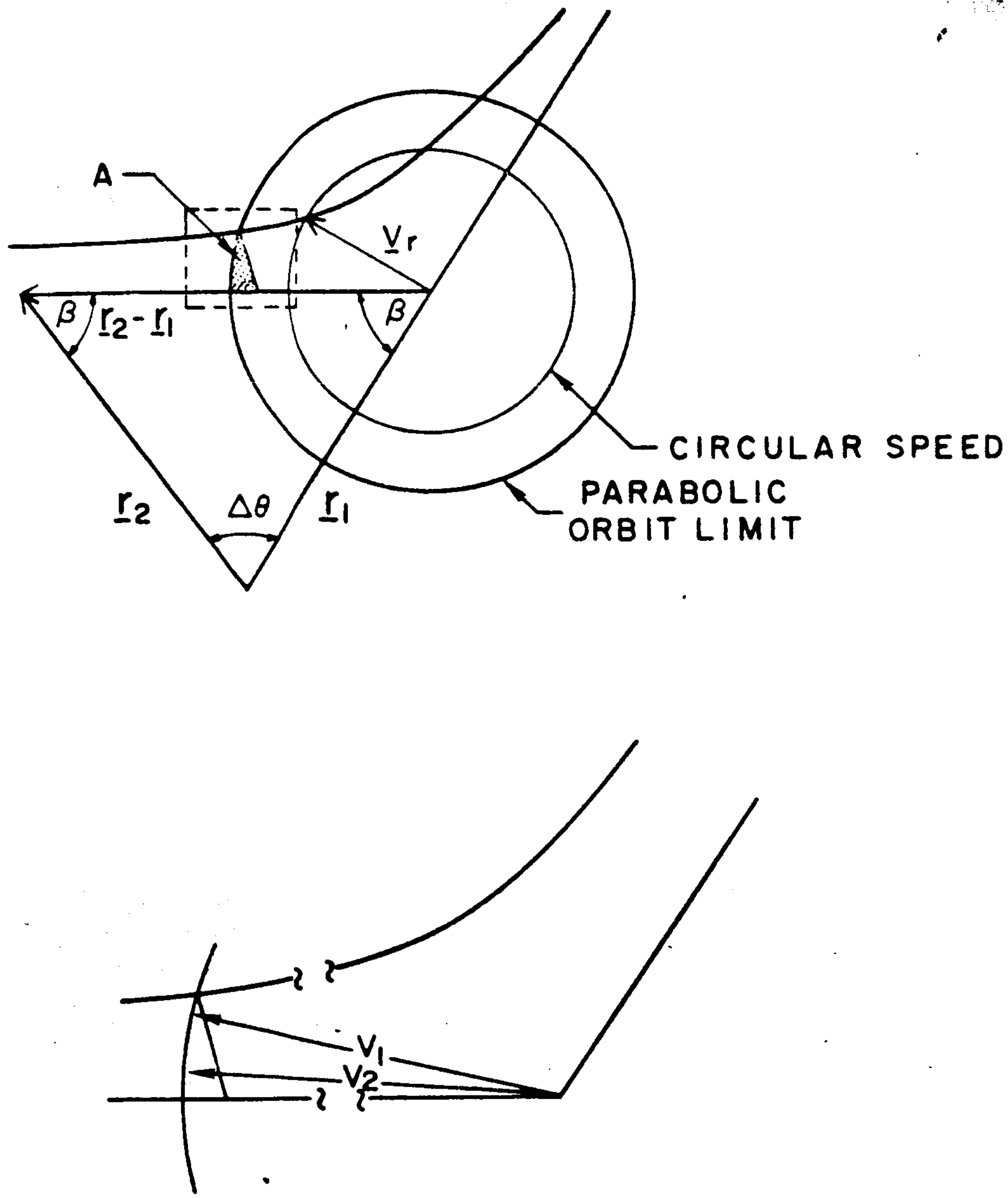


FIGURE 5. EXAMPLE OF HYPERBOLIC MINIMUM
CASE OF EQUAL RADII

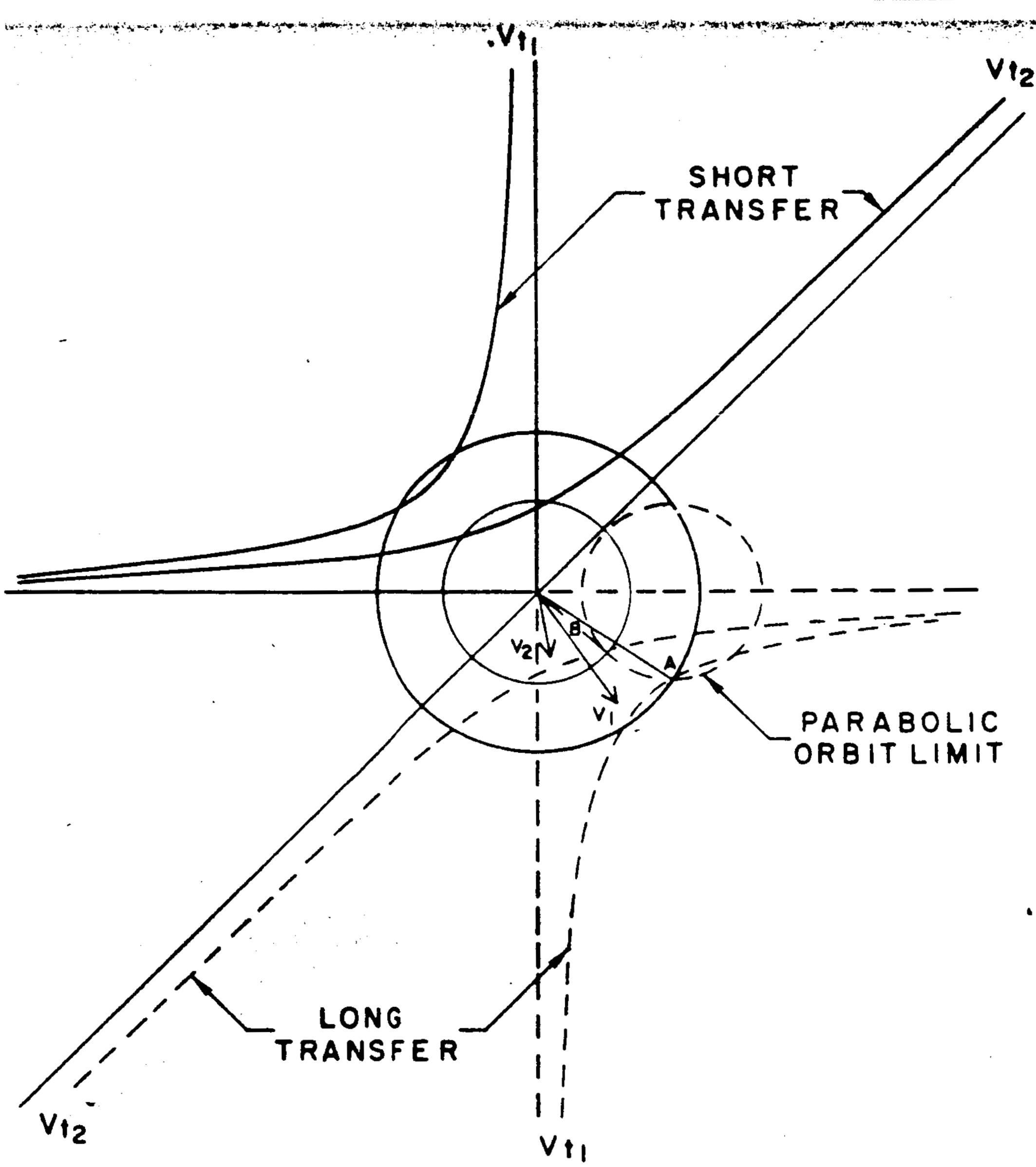


FIGURE 6. HYPERBOLIC MINIMUM TRANSFER — CASE OF NON-EQUAL RADII

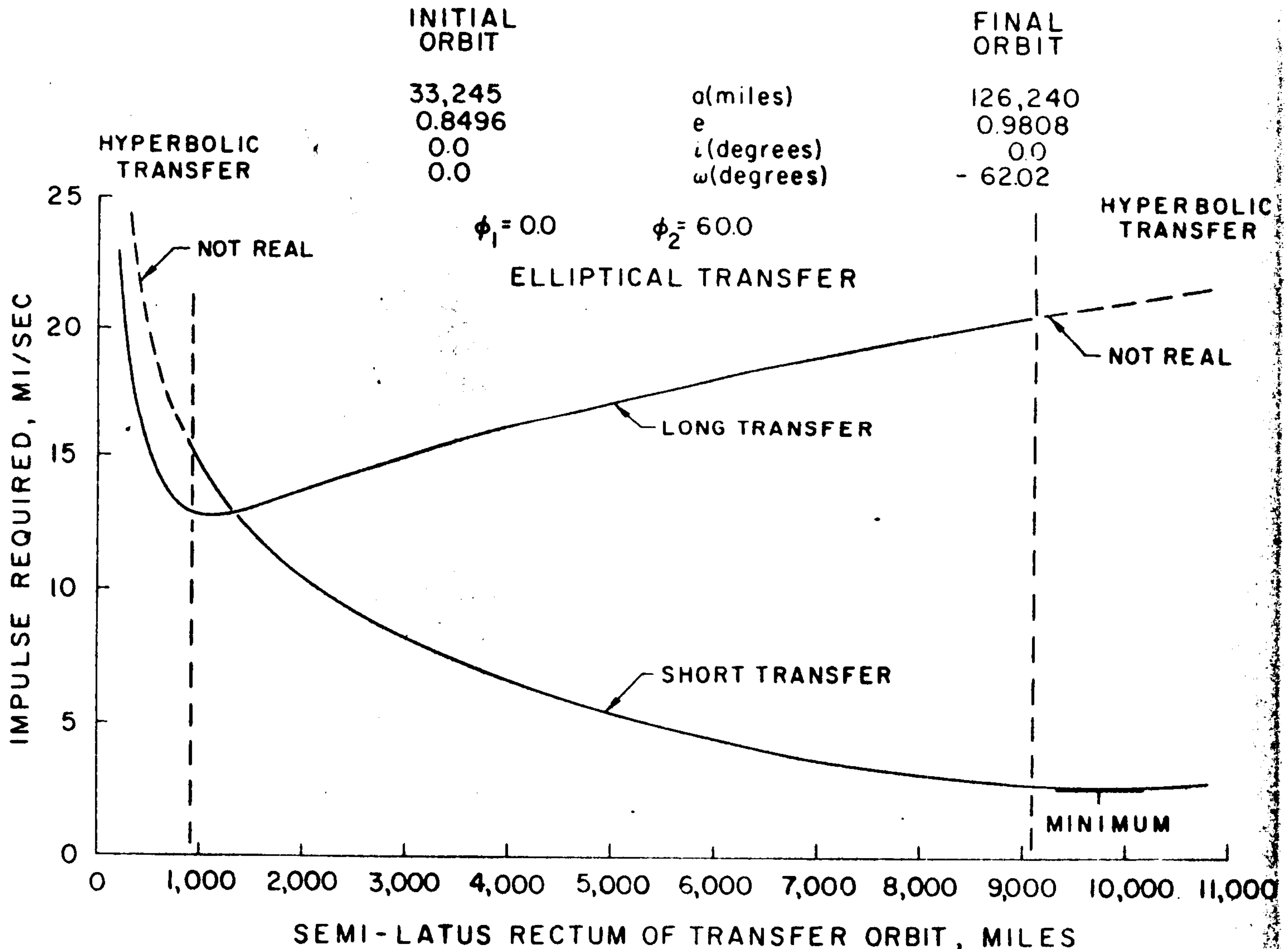


FIGURE 7. EXAMPLE OF HYPERBOLIC MINIMUM

VIII. ANALYSIS OF IMPULSE FUNCTION

The location of these peculiarities in the impulse function prompted an analytic search into the equations that describe the impulse problem. New analytic boundaries, different from the parabolic orbit limits, were sought for the minima. For fixed terminals (once again it should be pointed out that this is a restricted case of the more general problem of <u>optimizing</u> between any points on elliptical orbits), the impulse function is only dependent on p, the semi-latus rectum of the transfer orbit. This impulse function, defined by equation (21), has an extremum at all points p where

$$\frac{\partial I}{\partial p} = \frac{\partial \Psi_1}{\partial p} + \frac{\partial \Psi_2}{\partial p} = 0 \tag{26}$$

In the analysis of the impulse function carried out here, only the short transfers are considered. It is shown in a subsequent section that the extension to include the long transfers is very simple.

Now

$$\begin{aligned} \Psi_1(p) &= \left[(\underline{V}_{t1}(p) - \underline{V}_1) \cdot (V_{t1}(p) - \underline{V}_1) \right]^{\frac{1}{2}} \\ &= \left[(\underline{v}(p) + z(p)\,\underline{U}_1 - \underline{V}_1) \cdot (\underline{v}(p) + z(p)\,\underline{U}_1 - \underline{V}_1) \right]^{\frac{1}{2}} \\ &= \left[f(p) \right]^{\frac{1}{2}} \end{aligned} \tag{27}$$

where

$$\begin{aligned} f(p) &= \underline{v}(p) \cdot \underline{v}(p) + z^2(p) + \underline{V}_1 \cdot \underline{V}_1 - 2z(p)\,\underline{V}_1 \cdot \underline{U}_1 \\ &\quad - 2\underline{V}_1 \cdot \underline{v}(p) + 2z(p)\,\underline{v}(p) \cdot \underline{U}_1 \\ &= Ap + 2Bp^{\frac{1}{2}} + G - 2Cp^{-\frac{1}{2}} - Dp^{-1} \end{aligned} \tag{28}$$

where the coefficients are given in Table 1.

Similarly,

$$\Psi_2(p) = \left[g(p)\right]^{\frac{1}{2}} \tag{29}$$

where

$$g(p) = Ap + 2Ep^{\frac{1}{2}} + H - 2Fp^{-\frac{1}{2}} - Dp^{-1} \tag{30}$$

where the new coefficients are also given in Table 1.

Then, in order for impulse to be an extremum,

$$\frac{\partial \Psi_1}{\partial p} + \frac{\partial \Psi_2}{\partial p} = \frac{1}{2\Psi_1}\frac{\partial f}{\partial p} + \frac{1}{2\Psi_2}\frac{\partial g}{\partial p} = 0 \tag{31}$$

$$\Longrightarrow \frac{\Psi_1 \ (p)}{\Psi_2 \ (p)} = - \frac{\frac{\partial f}{\partial p}}{\frac{\partial g}{\partial p}} \tag{32}$$

Since $\Psi_1(p)$ and $\Psi_2(p)$ are always positive, it is easy to see from equation (32) that $\frac{\partial f}{\partial p}$ and $\frac{\partial g}{\partial p}$ must be of different sign before an extremum can occur in the impulse function. This important fact permits the identification of the extraneous roots in the eighth-order polynomial that will be derived.

Then

$$\frac{\partial f}{\partial p} = A + Bp^{-\frac{1}{2}} + Cp^{-3/2} + Dp^{-2} \tag{33}$$

and

$$\frac{\partial g}{\partial p} = A + Ep^{-\frac{1}{2}} + Fp^{-3/2} + Dp^{-2} \tag{34}$$

Before a meaningful expression can be worked out for the extrema in the impulse, equation (32) must be squared. Then the necessary expression becomes

$$\frac{f(p)}{g(p)} = \frac{\left(\frac{\partial f}{\partial p}\right)^2}{\left(\frac{\partial g}{\partial p}\right)^2}$$

or

$$f(p)\left(\frac{\partial g}{\partial p}\right)^2 - g(p)\left(\frac{\partial f}{\partial p}\right)^2 = 0 \qquad (35)$$

When this equation is multiplied out using equations (27), (28), (33), and (34), together with the substitution

$$s = p^{\frac{1}{2}} \qquad (36)$$

the necessary condition for an extremum becomes

$$\phi_1 s^8 + \phi_2 s^7 + \phi_3 s^6 + \phi_4 s^5 + \phi_5 s^4 + \phi_6 s^3 + \phi_7 s^2 + \phi_8 s + \phi_9 = 0 \qquad (37)$$

where the coefficients ϕ_i, i = 1-8 are given in Table 2. The real roots of this eighth-order polynomial must include all the values of p for which the impulse is an extremum.

The squaring process introduced in equation (35) added some extraneous roots to the octic—roots which do not correspond to extrema in I(p). These can be identified by factoring equation (35) as the difference of two squares.

$$f(p)\left(\frac{\partial g}{\partial p}\right)^2 - g(p)\left(\frac{\partial f}{\partial p}\right)^2 = 0$$

$$\Longrightarrow \left(\Psi_1(p)\frac{\partial g}{\partial p} + \Psi_2(p)\frac{\partial f}{\partial p}\right)\left(\Psi_1(p)\frac{\partial g}{\partial p} - \Psi_2(p)\frac{\partial f}{\partial p}\right) = 0 \qquad (38)$$

Since $\frac{\partial f}{\partial p}$ and $\frac{\partial g}{\partial p}$ must be of different sign, only those real values of p which are roots of

$$\Psi_1(p)\frac{\partial g}{\partial p} + \Psi_2(p)\frac{\partial f}{\partial p} = 0 \qquad (39)$$

are true extrema of I(p). It is easily shown (See Appendix 4) that the equation

$$\Psi_1(p)\,\frac{\partial g}{\partial p} - \Psi_2(p)\,\frac{\partial f}{\partial p} = 0 \qquad (40)$$

contains the extraneous roots of the octic and refers to extrema in another function, $I^*(p)$. Then

$$I^*(p) = \Psi_1(p) - \Psi_2(p) \qquad (41)$$

Inquiries into the nature of this octic suggest that four of these roots refer to extrema in $I^*(p)$. Although no general proof has been made, if this fact were true for all configurations, then there could be no more than two minima on either transfer branch. This would greatly simplify the application of the contour mapping approach.

In Mr. Altman's paper, he identifies an eighth-order polynomial, the roots of which refer to minima in the case of two-impulse orbital transfer between coplanar orbits. Equation (37) extends the analysis, using different techniques, both to include inclined orbits <u>and</u> to identify those roots of the equation that are extraneous and do not refer to minima in the impulse function.

IX. THE BOUNDARIES ON MINIMA

Since a necessary condition for the existence of an extremum in the impulse function is that $\frac{\partial f}{\partial p}$ and $\frac{\partial g}{\partial p}$ be of different sign, analyses were next directed to determine the values for p for which they could be of different sign.

From equations (33) and (34),

$$\lim_{p \to \infty} \frac{\partial f}{\partial p} = \lim_{p \to \infty} \frac{\partial g}{\partial p} = A \tag{42}$$

where

$$A = \frac{\mu \left| \underline{r}_2 - \underline{r}_1 \right|^2}{\left| \underline{r}_1 \times \underline{r}_2 \right|^2} > 0 \tag{43}$$

and

$$\lim_{p \to 0^+} \frac{\partial f}{\partial p} = \lim_{p \to 0^+} \frac{\partial g}{\partial p} = -\infty \tag{44}$$

because

$$D = -\mu \tan^2 \frac{\Delta\theta}{2} \tag{45}$$

Since for p both very small and very large, $\frac{\partial f}{\partial p}$ and $\frac{\partial g}{\partial p}$ have the same sign, we know that the region in which $\frac{\partial f}{\partial p}$ and $\frac{\partial g}{\partial p}$ are of different sign may definitely be bounded. The boundaries in which <u>all</u> minima in the impulse function (on short transfer side) must lie are given by the least positive value of p and the greatest positive value of p at which either

$$\frac{\partial f}{\partial p} = 0 \quad \text{or} \quad \frac{\partial g}{\partial p} = 0 \tag{46}$$

Since, for $s = p^{\frac{1}{2}}$, $\frac{\partial f}{\partial p} = 0$ where

$$As^4 + Bs^3 + Cs + D = 0 \tag{47}$$

Similarly, $\frac{\partial g}{\partial p} = 0$ where

$$As^4 + Es^3 + Fs + D = 0 \tag{48}$$

These values for p that bound the minima can be readily obtained. It is shown in a subsequent section that these equations also give the intervals for the long transfer. Thus, definite, analytic boundaries on the possible range of the impulse minima have been ascertained.

X. THE INTERVALS

Since both $\frac{\partial f}{\partial p}$ and $\frac{\partial g}{\partial p}$ have negative values for p very small and positive values for p very large, both expressions (47 and 48) must have an odd number of positive real roots. Each of these quartic equations may have either one or three real positive roots. Regardless how many of these roots each of these quartics has, all possible combinations of the roots can be studied by investigating two types of intervals in which $\frac{\partial f}{\partial p}$ and $\frac{\partial g}{\partial p}$ may be of different sign.

Type A: 1) $\frac{\partial f}{\partial p}$ and $\frac{\partial g}{\partial p}$ of different sign in $[a, b]$

$$2)\ \left(\frac{\partial f}{\partial p}\right)_{p=a} = 0;\ \left(\frac{\partial g}{\partial p}\right)_{p=b} = 0 \tag{49}$$

Type B: 1) $\frac{\partial f}{\partial p}$ and $\frac{\partial g}{\partial p}$ of different sign in $[a, b]$

$$2)\ \left(\frac{\partial f}{\partial p}\right)_{p=a} = 0;\ \left(\frac{\partial f}{\partial p}\right)_{p=a} = 0 \tag{50}$$

It is important to note that if, in equation (49), $\frac{\partial f}{\partial p}$ and $\frac{\partial g}{\partial p}$ are zero at opposite ends of the interval from those given, the problem is not really changed. Similarly, if in equation (50) it is $\frac{\partial g}{\partial p}$ which is zero at both ends, the analysis of the types of intervals still holds. In type A each of the functions is zero at one end of the interval; in type B, one function is zero at both ends of the interval. The two types of intervals are illustrated in Figure 8.

These intervals are divided into two types because the number of minima possible in a given interval is determined by its type. Define two functions $\sigma(p)$ and $\tau(p)$ as

$$\sigma(p) = -\frac{\frac{\partial f}{\partial i}}{\frac{\partial i}{\partial p}} \tag{51}$$

$$\tau(p) = \frac{\Psi_1(p)}{\Psi_2(p)} \tag{52}$$

Obviously, an extremum in the impulse function occurs for all p at which $\sigma(p) = \tau(p)$.

Consider an interval of Type A. τ is monotonic increasing and positive for all p in [a, b]. Also, note that

$$\sigma(a) = 0 \tag{53}$$

and

$$\lim_{p \to b} \sigma(p) = \infty \tag{54}$$

From Figure 8, it is clear σ and τ <u>must</u> intersect at least one time (producing one extremum) in that interval. If they are equal more than once, they must intersect an odd number of times.

Consider next an interval of Type B. Once again τ is monotonic increasing and positive for all p in [a, b]. Here, though

$$\sigma(a) = 0 \tag{55}$$

and

$$\sigma(b) = 0 \tag{56}$$

while for all p in [a, b], $\sigma(p) > 0$. It is evident from Figure 8 that σ and τ must intersect an even number of times in intervals of this type.

All possible permutations of the roots of these quartics can be manipulated to reduce the problem to an analysis of these intervals. Most frequently, both $\frac{\partial f}{\partial p}$ and $\frac{\partial g}{\partial p}$ have one real, positive root and produce an interval of type A in which σ and τ intersect one time. It is also true that, for the majority of the cases, the first and last real positive roots of the two quartics will limit the search for the minimum impulse to elliptical transfer orbits. These analyses do, however, explain the existence of the two peculiarities located earlier.

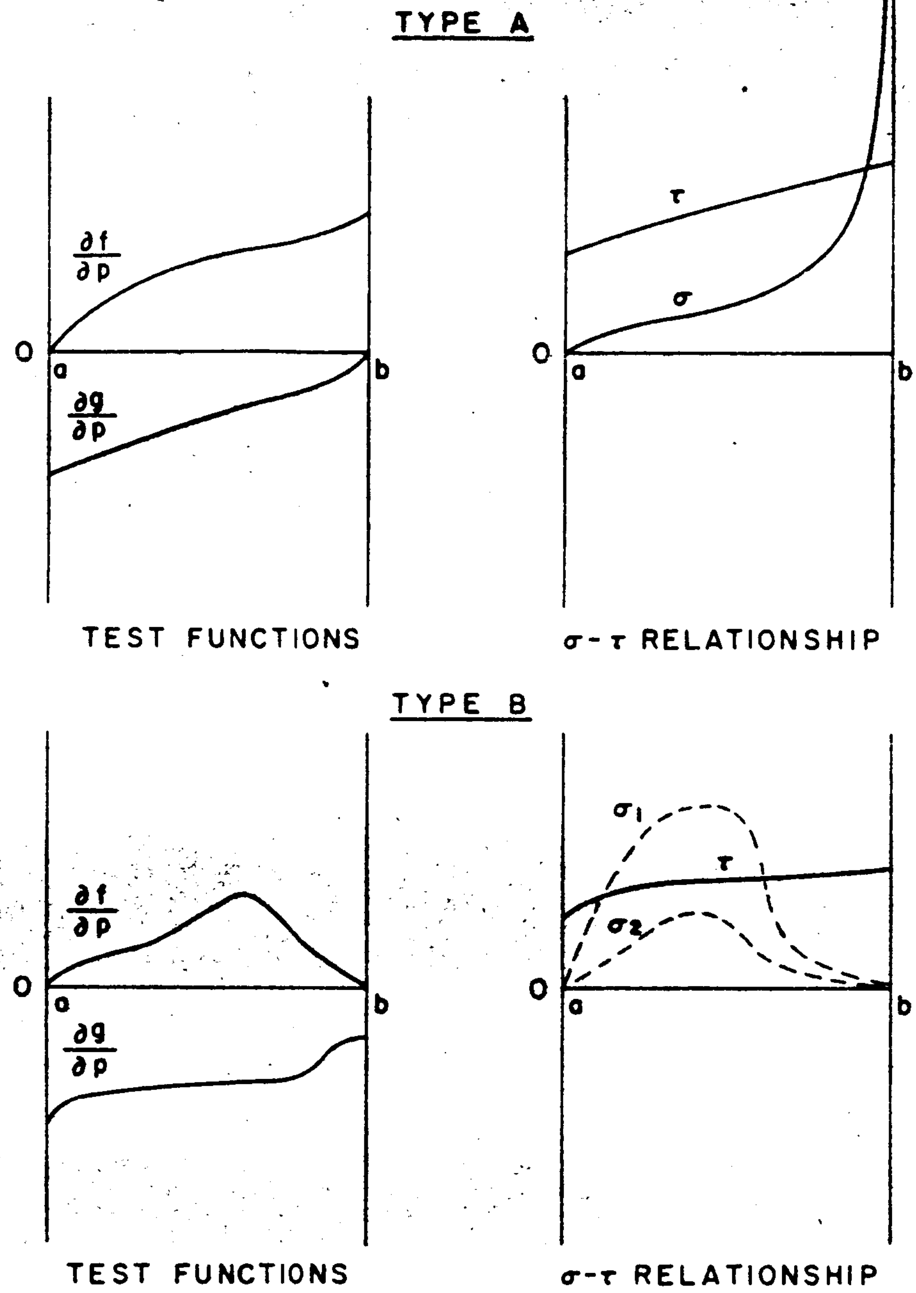


FIGURE 8 ILLUSTRATION OF INTERVALS

XI. LONG AND SHORT TRANSFER

For nearly all the equations derived in the preceding sections, it was assumed that the two-impulse orbital transfer was accomplished with a true anomaly interval in the transfer orbit of less than 180°--short transfer. The symmetry of the problem makes extension to include the long transfers very simple. To obtain the absolute minimum impulse, the two are then compared.

Because of the symmetry (See Appendix 5 for detailed derivation of long transfer equations) it can be shown that the real, negative roots of equations (33) and (34) determine intervals on the long transfer side that may produce minima. Similarly, it is the real, negative roots of the general octic (Equation 37) that appear within those specified intervals that determine values of p for which the long transfer may be an extremum.

This implies that all the analysis can be conducted by examining three equations--two quartic and one octic. The real roots of these equations--positive for short transfer and negative for long transfer--define all the intervals in which the extrema may exist and then locate the values of p at which extrema actually occur.

XII. MODIFICATION OF COMPUTER PROGRAM

In Mr. McCue's work, he optimizes the transfer between two elliptical orbits (not fixed terminals) by means of a contour mapping routine in the ϕ_1-ϕ_2 space that connects transfers of equal impulse requirement. In locating the minimum impulse for a particular ϕ_1-ϕ_2 configuration, he conducts a numerical search minimization along p that is confined within the parabolic orbit limits. The results of the analytic investigations presented here have been incorporated into the program to remove the limitations.

The solution of the two quartics (Equations 33 and 34) gives the intervals to which the numerical search for minima may be restricted. In nearly all cases a search within these limits will require fewer iterations than one conducted with the old arbitrary limits. If the steepest descent program should—in rare instances—converge on two different values within the same interval, the octic equation may be solved and its real roots compared to the intervals of possible minima. For nearly all ϕ_1-ϕ_2 configurations, this process will require no more computer time than it did before the limitations were removed.

TABLE 1

COEFFICIENTS OF TEST FUNCTIONS

$$A = \mu \frac{|\underline{r}_2 - \underline{r}_1|^2}{|\underline{r}_1 \times \underline{r}_2|^2}$$

$$B = - \frac{\mu}{(p_1)^{\frac{1}{2}}|\underline{r}_1 \times \underline{r}_2|} \left[(\underline{r}_2 - \underline{r}_1) \cdot (\underline{W}_1 \times (\underline{e}_1 + \underline{U}_1))\right]$$

$$C = \frac{\mu \tan \frac{\Delta\theta}{2}}{(p_1)^{\frac{1}{2}}} \left[U_1 \cdot (\underline{W}_1 \times \underline{e}_1)\right]$$

$$D = - \mu \tan^2 \frac{\Delta\theta}{2}$$

$$E = \frac{-\mu}{(p_2)^{\frac{1}{2}}|\underline{r}_1 \times \underline{r}_2|} \left[(\underline{r}_2 - \underline{r}_1) \cdot (\underline{W}_2 \times (\underline{e}_2 + \underline{U}_2))\right]$$

$$F = \frac{-\mu \tan \frac{\Delta\theta}{2}}{(p_2)^{\frac{1}{2}}} \left[\underline{U}_2 \cdot (\underline{W}_2 \times \underline{e}_2)\right]$$

$$G = \frac{\mu}{p_1}\left[\underline{W}_1 \times (\underline{e}_1 + \underline{U}_1)\right]^2 + \frac{2\mu \tan \frac{\Delta\theta}{2}}{|\underline{r}_1 \times \underline{r}_2|}\left[\underline{U}_1 \cdot (\underline{r}_2 - \underline{r}_1)\right]$$

$$H = \frac{\mu}{p_2}\left[\underline{W}_2 \times (\underline{e}_2 + \underline{U}_2)\right]^2 - \frac{2\mu \tan \frac{\Delta\theta}{2}}{|\underline{r}_1 \times \underline{r}_2|}\left[\underline{U}_2 \cdot (\underline{r}_2 - \underline{r}_1)\right]$$

XIII. SUMMARY

New analytical approaches to the two-impulse orbital transfer problem are developed in this paper. This development precipitated the discovery of both the hyperbolic minimum and the double minimum in the minimum velocity increment solution between points on elliptical orbits. Further analyses produced an eighth-order polynomial--applicable even for inclined orbits--whose roots contain all possible extrema in the impulse function. Next test functions were located that placed bounds on the regions in which these extrema could exist and identified those roots of the octic that were extraneous. The explanation of these extraneous roots--not corresponding to minima in the impulse function--was given.

All these results have been used to modify an earlier computer program. It is now possible to locate not only the absolute minimum two-impulse transfer between fixed terminals for any elliptical orbit pair, but also the absolute optimum transfer between any end points on those orbits.

TABLE 2

COEFFICIENTS OF EIGHTH-ORDER POLYNOMIAL

$$\phi_1 = A^2(G - H) + A(E^2 - B^2)$$

$$\phi_2 = A^2(4F - 4C) + A(2EG - 2BH) + 2E^2B - 2EB^2$$

$$\phi_3 = A(8BF - 8EC + 2EF - 2BC) + E^2G - HB^2$$

$$\phi_4 = A(4BD - 4ED + 2FG - 2CH) + 4BEF - 2CE^2 - 4BCE + 2FB^2$$

$$\phi_5 = D(2AG - 2HA - E^2 + B^2) + A(F^2 - C^2) - 2BCH + 2GEF$$

$$\phi_6 = D(4FA - 4AC + 2EG - 2BH) + 4FBC - 4CEF + 2BF^2 - 2EC^2$$

$$\phi_7 = D(8BF - 8EC + 2BC - 2EF) + F^2G - C^2H$$

$$\phi_8 = D^2(4B - 4E) + D(2FB - 2HC) + 2FC^2 - 2CF^2$$

$$\phi_9 = D^2(G - H) + D(C^2 - F^2)$$

REFERENCES

1. Smart, W. M., Celestial Mechanics, Longmans, Green and Company, London, 1953.

2. Bender, D. F., "Optimum Coplanar Two-Impulse Transfer Between Elliptic Orbits", Journal of Aerospace Engineering, Vol. 21, October 1962, p. 44-52.

3. Altman, S. P. and Pistiner, J. S., "Minimum Velocity Increment Solution for Two-Impulse Coplanar Orbital Transfer", AIAA Journal, Vol. 1, number 2, February 1963, pp. 435-442.

4. Vargo, L. G., "Optimal Transfer Between Coplanar Terminals in a Gravitational Field," Advances in Astronautical Sciences, Plenum Press, New York, 1958, Vol. 3, pp. 20-1 to 20-9.

5. McCue, G. A., "Optimum Two-Impulse Orbital Transfer and Rendezvous Between Inclined Elliptical Orbits", SID 62-1400, North American Aviation, Inc., November 20, 1962, (To be published in AIAA Journal).

6. Herget, Paul, "The Computation of Orbits", Published privately by author, Ann Arbor, Michigan, 1948, p. 30

7. Bell, H. W., Unpublished private communication, 1963.

8. Arenstorf, Richard A., "Solution of Two-Body Problem in Complex Form", Unpublished written analysis, 1960.

APPENDIX 1

Derivation of Transfer Velocity Expressions:

These equations appear in reference (5) and were originally derived by Mr. H. W. Bell.

Begin with the vector expressions for the transfer orbit velocities,

$$\underline{V}_{t1}(p) = \left(\frac{\mu}{p}\right)^{\frac{1}{2}} \underline{W}_t \times (\underline{e}_t + \underline{U}_1) \tag{A1-1}$$

$$\underline{V}_{t2}(p) = \left(\frac{\mu}{p}\right)^{\frac{1}{2}} W_t \times (\underline{e}_t + \underline{U}_2) \tag{A1-2}$$

where $\underline{e}_t$ is a vector, not necessarily explicitly defined, that has the magnitude of the eccentricity of the transfer orbit and is in the direction of its perigee.

Then for all transfer orbits that include $\underline{r}_1$ and $\underline{r}_2$,

$$|\underline{r}_1| = \frac{p}{1 + \underline{e}_t \cdot \underline{U}_1} \tag{A1-3}$$

$$|\underline{r}_2| = \frac{p}{1 + \underline{e}_t \cdot \underline{U}_2} \tag{A1-4}$$

and by algebraic manipulation

$$\underline{e}_t \cdot \underline{U}_1 = \frac{p}{|\underline{r}_1|} - 1 \tag{A1-5}$$

$$\underline{e}_t \cdot \underline{U}_2 = \frac{p}{|\underline{r}_2|} - 1 \tag{A1-6}$$

Then multiplying these equations by $\underline{U}_2$ and $\underline{U}_1$ gives

$$(\underline{e}_t \cdot \underline{U}_1)\,\underline{U}_2 = \underline{U}_2\left(\frac{p}{|\underline{r}_1|} - 1\right) \tag{A1-7}$$

and

$$(\underline{e}_t \cdot \underline{U}_2)\,\underline{U}_1 = \underline{U}_1\left(\frac{p}{|\underline{r}_2|} - 1\right) \tag{A1-8}$$

Then according to vector identities

$$(\underline{U}_1 \times \underline{U}_2) \times \underline{e}_t = -\underline{e}_t \times (\underline{U}_1 \times \underline{U}_2)$$

$$= -\,\underline{U}_1\,(U_2 \cdot \underline{e}_t) + U_2\,(\underline{U}_1 \cdot \underline{e}_t)$$

$$= \underline{U}_2\left(\frac{p}{|\underline{r}_1|} - 1\right) - \underline{U}_1\left(\frac{p}{|\underline{r}_2|} - 1\right) \tag{A1-9}$$

Notice that

$$\underline{W}_t \times (\underline{e}_t + \underline{U}_1) = \frac{(\underline{U}_1 \times \underline{U}_2)}{|\underline{U}_1 \times \underline{U}_2|} \times (\underline{e}_t + \underline{U}_1)$$

$$= \frac{(\underline{U}_1 \times \underline{U}_2) \times \underline{e}_t}{\underline{U}_1 \times \underline{U}_2} + \frac{(\underline{U}_1 \times \underline{U}_2) \times \underline{U}_1}{U_1 \times \underline{U}_2}$$

$$= \frac{\underline{U}_2\left(\frac{p}{|\underline{r}_1|} - 1\right) - \underline{U}_1\left(\frac{p}{|\underline{r}_2|} - 1\right) + \underline{U}_2\,(\underline{U}_1 \cdot \underline{U}_1) - \underline{U}_1\,(\underline{U}_1 \cdot \underline{U}_2)}{|\underline{U}_1 \times \underline{U}_2|} \tag{A1-10}$$

Similarly,

$$\underline{W}_t \times (\underline{e}_t + \underline{U}_2) = \frac{\underline{U}_2\left(\frac{p}{|\underline{r}_1|} - 1\right) - \underline{U}_1\left(\frac{p}{|\underline{r}_2|} - 1\right) - \underline{U}_1\,(\underline{U}_2 \cdot \underline{U}_2) + \underline{U}_2\,(\underline{U}_1 \cdot \underline{U}_2)}{\underline{U}_1 \times \underline{U}_2} \tag{A1-11}$$

Then from Equations (A1-1), (A1-2), (A1-10), and (A1-11),

$$V_{t1}(p) = \left(\frac{\mu}{p}\right)^{\frac{1}{2}} \frac{1}{|\underline{U}_1 \times \underline{U}_2|}\left[p\left(\frac{\underline{U}_2}{|\underline{r}_1|} - \frac{\underline{U}_1}{|\underline{r}_2|}\right) + (1 - \underline{U}_1 \cdot \underline{U}_2)\,\underline{U}_1\right] \tag{A1-12}$$

and

$$V_{t2}(p) = \left(\frac{\mu}{p}\right)^{\frac{1}{2}} \frac{1}{|\underline{U}_1 \times \underline{U}_2|} \left[p \left(\frac{\underline{U}_2}{|\underline{r}_1|} - \frac{\underline{U}_1}{|\underline{r}_2|} \right) - (1 - \underline{U}_1 \cdot \underline{U}_2)\, \underline{U}_2 \right] \tag{A1-13}$$

note that

$$U_1 \cdot U_2 = \cos \Delta\theta \tag{A1-14}$$

and

$$|\underline{U}_1 \times \underline{U}_2| = |U_1|\ |U_2| \sin \Delta\theta = \sin \Delta\theta \tag{A1-15}$$

Furthermore,

$$\frac{p}{|\underline{U}_1 \times \underline{U}_2|} \left[\frac{\underline{U}_2}{|\underline{r}_1|} - \frac{\underline{U}_1}{|\underline{r}_2|} \right] = \frac{p}{|\underline{U}_1 \times \underline{U}_2|} \left[\frac{\underline{U}_2\ \underline{r}_2 - \underline{U}_1\ \underline{r}_1}{|\underline{r}_1|\ |\underline{r}_2|} \right]$$

$$= \frac{p\,(\underline{r}_2 - \underline{r}_1)}{|\underline{r}_1 \times \underline{r}_2|} \tag{A1-16}$$

therefore,

$$\underline{V}_{t1}(p) = \left(\frac{\mu}{p}\right)^{\frac{1}{2}} \left[\frac{p\,(\underline{r}_2 - \underline{r}_1)}{\underline{r}_1 \times \underline{r}_2} + \frac{1 - \cos \Delta\theta}{\sin \Delta\theta}\, \underline{U}_1 \right] \tag{A1-17}$$

But

$$\frac{1 - \cos \Delta\theta}{\sin \Delta\theta} = \tan \frac{\Delta\theta}{2}\, * \tag{A1-18}$$

which implies that

$$\underline{V}_{t1}(p) = \frac{\mu}{p}^{\frac{1}{2}} \left[\frac{p\,(\underline{r}_2 - \underline{r}_1)}{\underline{r}_1 \times \underline{r}_2} + \tan \frac{\Delta\theta}{2}\, \underline{U}_1 \right]$$

$$= \underline{v} + z\underline{U}_1 \tag{A1-19}$$

---

* $\sin \Delta\theta = \underline{W}_t \cdot (\underline{U}_1 \times \underline{U}_2)$—this implies that angle from $\underline{U}_1$ to $\underline{U}_2$ is less than 180°.

Similarly, from (A1-13),

$$\underline{V}_{t2}(p) = \underline{v} - z\,\underline{U}_2 \qquad \text{(A1-20)}$$

where

$$\underline{v} = \frac{(\mu p)^{\frac{1}{2}}\,(\underline{r}_2 - \underline{r}_1)}{|\underline{r}_1 \times \underline{r}_2|} \qquad \text{(A1-21)}$$

and

$$z = \left(\frac{\mu}{p}\right)^{\frac{1}{2}}\left(\tan\frac{\Delta\theta}{2}\right) \qquad \text{(A1-22)}$$

APPENDIX 2

On the Criteria for Existence of Double Minimum in Case of Equal Radii:

After the existence of the double minimum was first established, the next research was directed toward finding the necessary and sufficient conditions for this existence. In the equal radii case, the answer was more or less obtainable.

If, for any $\underline{r}_1 - \underline{r}_2$ configuration where $|\underline{r}_1| = |\underline{r}_2|$, the transfer velocity (See Figure 2) hyperbola is considered to be symmetrical about an x-axis of a rectangular Cartesian coordinate system, the analytic equation of this hyperbola can be derived. From the general expression,

$$\frac{x^2}{a^2} - \frac{y^2}{b^2} = 1 \tag{A2-1}$$

and the knowledge that

$$\tan\left(45 - \frac{\Delta\theta}{4}\right) = \frac{a}{b} \tag{A2-2}$$

and the fact that there exists one point p_o on the hyperbola where

$$p_o = \left(\frac{\mu}{r}\right)^{\frac{1}{2}} \sin\Psi\, \underline{i} + \left(\frac{\mu}{r}\right) \cos\Psi\, \underline{j} \tag{A2-3}$$

$$\Psi = 45° + \frac{\Delta\theta}{4} \tag{A2-4}$$

the equation for this hyperbola is given as

$$r(y^2 \cos^2\Psi - x^2 \sin^2\Psi) = \mu \cos 2\Psi \tag{A2-5}$$

where

$$|\underline{r}_2| = |\underline{r}_1| = r \tag{A2-6}$$

$$a^2 = -\frac{\mu \cos 2\Psi}{r \sin^2\Psi} \tag{A2-7}$$

$$b^2 = -\frac{\mu \cos 2\Psi}{r \cos^2\Psi} \tag{A2-8}$$

The two elliptical orbits, initial and final, are defined by their velocity vectors at $\underline{r}_1$ and $\underline{r}_2$. The coordinates of these vectors in the same coordinate system as the hyperbola are readily found from the geometry and elementary celestial mechanics[1].

Suppose*

$$\underline{V}_1 = x_1 \underline{i} + y_1 \underline{j} \tag{A2-9}$$

$$\underline{V}_2 = x_2 \underline{i} + y_2 \underline{j} \tag{A2-10}$$

Then for every pair of elliptical orbits, two points in this system are determined. There is also a corresponding magnitude d_v, where

$$d_v = \left[(x_2 - x_1)^2 + (y_2 - y_1)^2 \right]^{\frac{1}{2}} \tag{A2-11}$$

In the Bell-Arenstorf diagram, it is important to remember that for the case where $\underline{r}_1 = \underline{r}_2$, impulse for a particular transfer is nothing more than the sum of the distances from $\underline{V}_1$ and $\underline{V}_2$ to the point on the hyperbola representing that transfer.

For any initial and final elliptical orbits, let the points defined by $\underline{V}_1$ and $\underline{V}_2$ in this coordinate system be considered as the foci for a family of confocal ellipses. Each scalar value k,

---

* The actual analytic expressions for these coordinates have been omitted because of their detail.

where $k \geq d_v$ (A2-12)

defines a member of this family. Let

$$k = s_1 + s_2 \tag{A2-13}$$

Then s_1 and s_2 are, for any particular ellipse, the distances from $\underline{V}_1$ and $\underline{V}_2$ to that ellipse—their sum must be a constant for any member of the family.

In terms of the Bell-Arenstorf diagram, each member of the family corresponds to a certain impulse value--as k grows larger, eventually a member of the family (See FIgure A2-1) intersects the hyperbola at a point of tangency. That point on the hyperbola, representing a particular transfer orbit, must be a relative minimum in impulse. This is easily seen if the next member of the family is considered—it intersects the hyperbola at two points, one on either side of the earlier point of tangency—it represents higher impulse. It is an easy intuitive extension to realize that every point at which a member of the family of ellipses is tangent to either branch of the hyperbola produces a relative extremum in either the long or short transfer.

In Figure A2-1, on the right, an example is given of a typical $\underline{V}_1 - \underline{V}_2$ configuration that produces one point of tangency--one relative minimum--on each of the two branches. The other example is a $\underline{V}_1 - \underline{V}_2$ configuration that produces a double minimum (see members 2 and 3 of the family) by having three points of tangency on one branch of the hyperbola. Note that if the family of ellipses has three members tangent to one branch of the hyperbola, there must exist some member of the family that intersects one branch of the hyperbola four times.

Since the analytic equation for this family of ellipses, using $s_1 + s_2 = k$ as the variable parameter, can be readily derived, the criteria for the existence of a double minimum can be simplified--for some member of the family, the fourth-order polynomial representing the intersection of that ellipse and the velocity hyperbola has four real roots of the same sign (on the same branch) if and only if a double minimum exists on that branch.

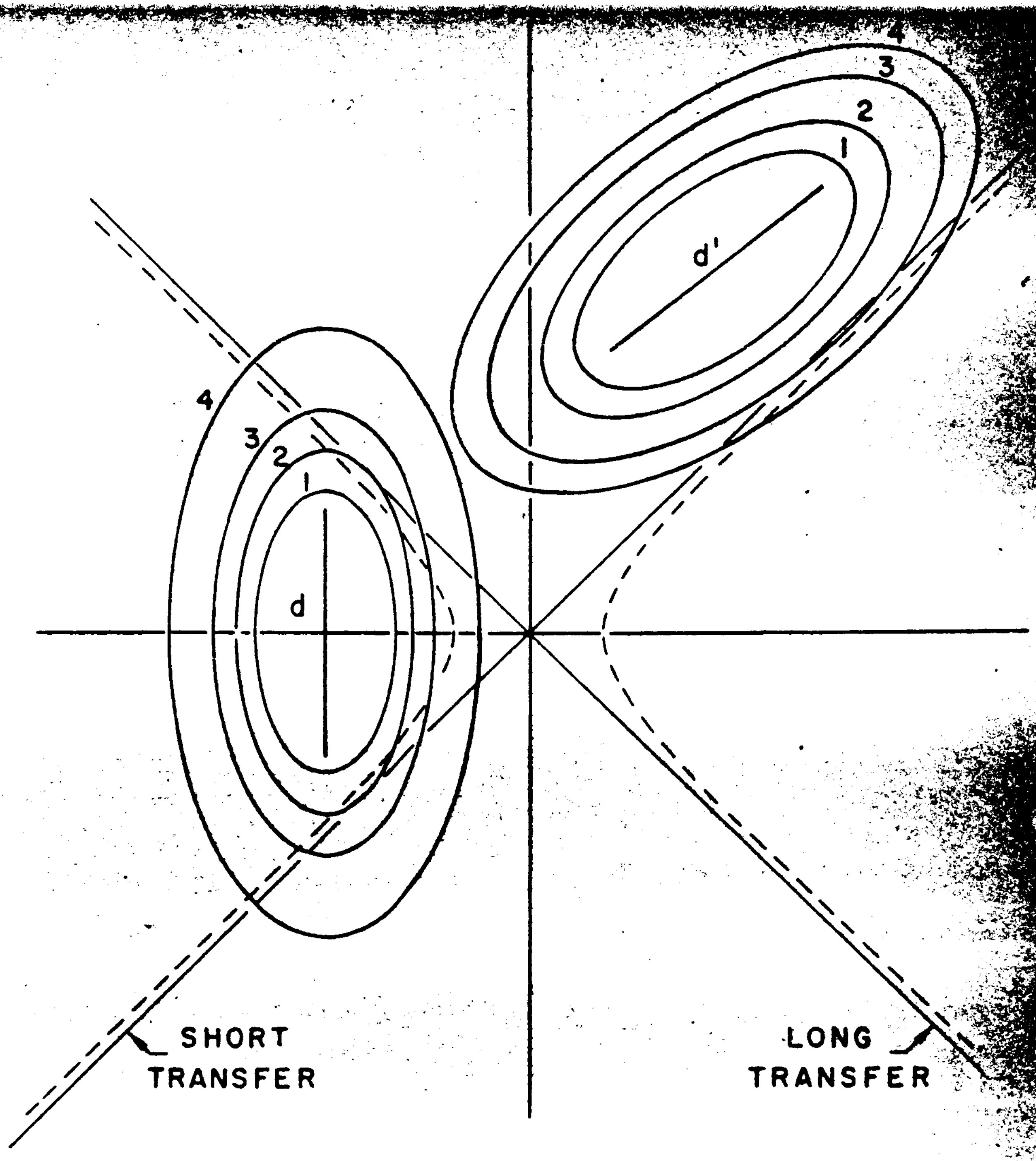


FIGURE A2-1 GEOMETRY OF DOUBLE MINIMUM
- IN EQUAL RADII CASE

APPENDIX 3

Method of Determining Corresponding Transfers on Velocity Hyperbolae:

In Figure 6, the hyperbola representing all possible transfer orbits that arrive at $\underline{r}_2$ and pass through $\underline{r}_1$ is transformed so that both of the hyperbolae have the same reference point and one common asymptote. It is next necessary to determine the corresponding points on the hyperbolae: that is, to find some way of relating—transfer orbit for transfer orbit—the point on one hyperbola with the point defining the same transfer orbit on the other hyperbola. Because of the high symmetry and so that Figure 6 may be used, only the long transfer has been considered here.

Then if

$$k = \frac{(\mu)^{\frac{1}{2}} \, |\underline{r}_2 - \underline{r}_1|}{|\underline{r}_1 \times \underline{r}_2|} \tag{A3-1}$$

and

$$h = \sqrt{\mu} \tan \frac{\Delta\theta}{2} \tag{A3-2}$$

and $\bar{m}$ is defined as a unit vector in the direction of $\underline{r}_2 - \underline{r}_1$.

$$V_{t1}(p) = -kp^{\frac{1}{2}} \, \underline{m} - hp^{-\frac{1}{2}} \, \underline{U}_1 \tag{A3-3}$$

$$V_{t2}(p) = -kp^{\frac{1}{2}} \, \underline{m} + hp^{-\frac{1}{2}} \, \underline{U}_2 \tag{A3-4}$$

It is obvious then that these two velocity hyperbolae have equal components in the m-direction. Furthermore, for each vector, the magnitude of the <u>other</u> component is the same. Thus the only significant difference in these functions is given by the different directions of $\underline{U}_1$ and $\underline{U}_2$.

For any real, positive value of p, the component of each of these transfer velocity vectors in the $\underline{r}_2 - \underline{r}_1$ direction is given by

$$\underline{V}_{t1} \text{ (comp } \underline{r}_2 - \underline{r}_1) = \underline{V}_{t2} \text{ (comp } \underline{r}_2 - \underline{r}_1) = - kp^{\frac{1}{2}} \qquad \text{(A3-5)}$$

Therefore for any p, there exists some point on the $r_2 - r_1$ axis that corresponds to this component. Suppose a circle of radius r^*, where

$$r^* = hp^{-\frac{1}{2}} \qquad \text{(A3-6)}$$

is circumscribed about that point as center. Then both the transfer velocity vectors for that particular transfer must end on that circle. Therefore, for any p, there exists a circle with center at a point on the $\underline{r}_2 - \underline{r}_1$ axis that will intersect the hyperbolae at the corresponding points.

It is an easy extension, therefore, to realize that the corresponding points on the hyperbolae are located by a family of circles with variable radii and centered on the $\underline{r}_2 - \underline{r}_1$ axis--the center and the radius being functions of the variable parameter p defining the transfer orbit.

APPENDIX 4

Identification of Extraneous Roots to Octic:

The eighth-order polynomial expression (Equation 37), whose roots contain the values of p at which the impulse function has a minimum, also has some roots that do not refer to impulse extrema.

This octic can be factored as the difference of two squares to produce Equation (38),

$$\left(\Psi_1(p)\ \frac{\partial g}{\partial p} + \Psi_2(p)\ \frac{\partial f}{\partial p}\right)\left(\Psi_1(p)\ \frac{\partial g}{\partial p} - \Psi_2(p)\ \frac{\partial f}{\partial p}\right) = 0 \qquad \text{(A4-1)}$$

It has already been shown that the equation

$$\left(\Psi_1(p)\ \frac{\partial g}{\partial p} + \Psi_2(p)\ \frac{\partial f}{\partial p}\right) = 0 \qquad \text{(A4-2)}$$

gives roots to the octic whose p values do correspond to extrema in the impulse function.

The other factor

$$\left(\Psi_1(p)\ \frac{\partial g}{\partial p} - \Psi_2(p)\ \frac{\partial f}{\partial p}\right) = 0 \qquad \text{(A4-3)}$$

produces roots to the octic that have no correspondence with impulse extrema.

Consider the function $I^*(p)$, where

$$I^*(p) = \Psi_1(p) - \Psi_2(p) \qquad \text{(A4-4)}$$

Then $I^*(p)$ has extrema at every value p where

$$\frac{\partial \Psi_1}{\partial p} - \frac{\partial \Psi_2}{\partial p} = \frac{1}{2\Psi_1}\ \frac{\partial f}{\partial p} - \frac{1}{2\Psi_2}\ \frac{\partial g}{\partial p} = 0 \qquad \text{(A4-5)}$$

This occurs whenever

$$\frac{\Psi_1(p)}{\Psi_2(p)} = \frac{\frac{\partial f}{\partial p}}{\frac{\partial g}{\partial p}}$$

or

$$\Psi_1(p) \frac{\partial g}{\partial p} - \Psi_2(p) \frac{\partial f}{\partial p} = 0 \qquad (A4-6)$$

This identifies the extraneous roots.

This function $I^*(p)$ corresponds to the <u>difference</u> between the magnitudes of the two velocity increments. Cases where this function has extrema for real values of p are not hard to locate. However, every real extremum for this function must lie in an interval in which $\frac{\partial f}{\partial p}$ and $\frac{\partial g}{\partial p}$ are the same sign. By using the interval technique described in section XI, these roots can be identified.

APPENDIX 5

Derivation of Long Transfer Equations:

For the "long transfer", the lower sign on the double-sign expressions for $\Psi_1(p)$ and $\Psi_2(p)$ (see equations (22) and 23)) is used. Thus

$$\begin{aligned}\Psi_1(p) &= \left|-\underline{V}_{t1}(p) - \underline{V}_1\right| = \left|\underline{V}_{t1}(p) + \underline{V}_1\right| \\ &= \left[(\underline{V}_{t1}(p) + \underline{V}_1) \cdot (\underline{V}_{t1}(p) + \underline{V}_1)\right]^{\frac{1}{2}} \\ &= \left[(\underline{v}(p) + z(p)\,\underline{U}_1 + \underline{V}_1) \cdot (\underline{v}(p) + z(p)\,\underline{U}_1 + \underline{V}_1)\right]^{\frac{1}{2}} \\ &= \left[h(p)\right]^{\frac{1}{2}} \end{aligned} \tag{A5-1}$$

where

$$\begin{aligned} h(p) &= \underline{v}(p) \cdot \underline{v}(p) + z^2(p) + \underline{V}_1 \cdot \underline{V}_1 + 2z(p)\,\underline{v}(p) \cdot \underline{U}_1 \\ &\quad + 2\underline{V}_1 \cdot \underline{v}(p) + 2z(p)\,\underline{V}_1 \cdot \underline{U}_1 \\ &= Ap - 2Bp^{\frac{1}{2}} + G + 2Cp^{-\frac{1}{2}} - Dp^{-1} \end{aligned} \tag{A5-2}$$

where the coefficients are given in Table 1.

Similarly,

$$\Psi_2(p) = \left[k(p)\right]^{\frac{1}{2}} \tag{A5-3}$$

where

$$k(p) = Ap - 2Ep^{\frac{1}{2}} + H + 2Fp^{-\frac{1}{2}} - Dp^{-1} \tag{A5-4}$$

and these coefficients are also found in Table 1.

Equation (32) is now replaced by the following criterion for the value of p at which the long transfer impulse has a minimum

$$\frac{\Psi_1(p)}{\Psi_2(p)} = -\frac{\frac{\partial h}{\partial p}}{\frac{\partial k}{\partial p}} \qquad \text{(A5-5)}$$

Now $\frac{\partial h}{\partial p}$ and $\frac{\partial k}{\partial p}$ must be of different sign in order for the long transfer impulse to have a minimum. Then

$$\frac{\partial h}{\partial p} = A - Bp^{-\frac{1}{2}} - Cp^{-3/2} + Dp^{-2} \qquad \text{(A5-6)}$$

and

$$\frac{\partial k}{\partial p} = A - Ep^{-\frac{1}{2}} - Fp^{-3/2} + Dp^{-2} \qquad \text{(A5-7)}$$

Then

$$\lim_{p \to \infty} \frac{\partial h}{\partial p} = \lim_{p \to \infty} \frac{\partial k}{\partial p} = A \qquad \text{(A5-8)}$$

and

$$\lim_{p \to 0^+} \frac{\partial k}{\partial p} = \lim_{p \to 0^+} \frac{\partial k}{\partial p} = -\infty \qquad \text{(A5-9)}$$

Thus by an analogy similar to the short transfer case, the minima must be in bounded intervals. These intervals can be found by analyzing the roots of two fourth-order polynomials

$$\frac{\partial h}{\partial p} = 0 \Longrightarrow As^4 - Bs^3 - Cs + D = 0 \qquad \text{(A5-10)}$$

where $s = p^{\frac{1}{2}}$

and

$$\frac{\partial k}{\partial p} = 0 \Longrightarrow As^4 - Es^3 - Fs + D = 0 \qquad \text{(A5-11)}$$

These fourth-order equations that produce bounds on the long transfer minima are very similar to those (see equations (47) and (48)) that provided the regions for the short transfer minima. In fact, if

$s = a$ is a real root of equation (A5-10), then $s = -a$ must be a root of equation (47). The same correspondence holds for the roots $\frac{\partial k}{\partial p} = 0$ and $\frac{\partial \bar{s}}{\partial p} = 0$.

Since the only values of p that are of practical interest in either case are for p real and positive, by simply analyzing the real roots of equations (47) and (48), both the long and short transfer intervals can be ascertained.

An eighth-order polynomial whose roots contain the impulse minima also exists for the long transfer. By squaring equation (A5-5), the necessary expression becomes

$$h(p)\left(\frac{\partial k}{\partial p}\right)^2 - k(p)\left(\frac{\partial h}{\partial p}\right)^2 = 0$$

If this expression is multiplied out in terms of equations (A5-3), (A5-4), (A5-6), and (A5-7), the necessary condition for the existence of an extremum in the long transfer impulse becomes

$$\phi_1 s^8 - \phi_2 s^7 + \phi_3 s^6 - \phi_4 s^5 + \phi_5 s^4 - \phi_6 s^3 + \phi_7 s^2 - \phi_8 s + \phi_9 = 0$$

where for $i = 1$-8, the ϕ_i coefficients are defined in Table 2.

Once again, if $s = a$ is a real root of equation (A5-13), then $s = -a$ is a real root of equation (37). Since the only values of p that are of practical interest are real and positive, both the long and short transfer extrema can be located in the analysis of the single octic. This symmetry would have been suggested by the Bell-Arenstorf technique.